



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

Walters, Charlotte L., Freeman, Robin, Collen, Alanna, Dietz, Christian, Fenton, M. Brock, Jones, Gareth, Obrist, Martin K, Puechmaille, Sébastien J., Sattler, Thomas, Siemers, Björn M., [Parsons, Stuart](#), & Jones, Kate E. (2012)

A continental-scale tool for acoustic identification of European bats.
Journal of Applied Ecology, 49(5), pp. 1064-1074.

This file was downloaded from: <https://eprints.qut.edu.au/79740/>

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

<https://doi.org/10.1111/j.1365-2664.2012.02182.x>

A continental-scale tool for acoustic identification of European bats

Charlotte Walters

Robin Feeman

Alana Collen

Christian Dietz

M. Brock fenton

Gareth Jones

Martin K. Obrist

Sebastien Puechmaille

Thomas Sattler

Bjorn M. Siemers

Stuart Parsons

Kate E. Jones

Abstract

Acoustic methods are used increasingly to survey and monitor bat populations. However, the use of acoustic methods at continental scales can be hampered by the lack of standardized and objective methods to identify all species recorded. This makes comparable continent-wide monitoring difficult, impeding progress towards developing biodiversity indicators, trans-boundary conservation programmes and monitoring species distribution changes.

Here we developed a continental-scale classifier for acoustic identification of bats, which can be used throughout Europe to ensure objective, consistent and comparable species identifications. We selected 1350 full-spectrum reference calls from a set of 15 858 calls of 34 European species, from EchoBank, a global echolocation call library. We assessed 24 call parameters to evaluate how well they distinguish between species and used the 12 most useful to train a hierarchy of ensembles of artificial neural networks to distinguish the echolocation calls of these bat species.

Calls are first classified to one of five call-type groups, with a median accuracy of 97.6%. The median species-level classification accuracy is 83.7%, providing robust classification for most European species, and an estimate of classification error for each species.

These classifiers were packaged into an online tool, iBatsID, which is freely available, enabling anyone to classify European calls in an objective and consistent way, allowing standardized acoustic identification across the continent.

Synthesis and applications. iBatsID is the first freely available and easily accessible continental-scale bat call classifier, providing the basis for standardized, continental acoustic bat monitoring in Europe. This method can provide key information to managers and conservation planners on distribution changes and changes in bat species activity through time.

Introduction

Bats are ideal candidates as indicators of habitat quality and climate change as they are globally distributed and provide essential ecosystem services (Jones et al. 2009). They also have traits making them particularly sensitive to human impacts, such as slow population growth rates (Jones & Maclarnon 2001) and temperature-sensitive hibernation behaviour (Jones et al. 2009). Survey and monitoring of bats is therefore not only important for assessing how bats are faring in response to a range of threats (Puechmaille et al. 2011a) but may also enable an understanding of the changing state of biodiversity in general.

The need to survey, monitor and protect bat populations has been recognized for some time, not least in Europe where bats are protected under the EUROBATS agreement (www.eurobats.org). Although standardized bat monitoring protocols for Europe have been proposed (Battersby 2010), these are yet to be adopted for any continent-wide monitoring programme. Continental-scale survey and monitoring is important in developing global biodiversity indicators (Pereira & Cooper 2006), for tracking distribution changes and for trans-boundary conservation relevant to large-scale species distributions (Poiani et al. 2000), particularly for migratory species.

As most bats are small, nocturnal and often difficult to catch, survey programmes that rely on visual encounters or captures of individuals (Robbins, Bystrak & Geissler 1986; Kery & Schmid 2005) are difficult to apply to bats, and unlikely to be efficient for surveying bat populations at large scales (Ochoa, O'Farrell & Miller 2000). Many bats use echolocation for orientation and prey detection (Pierce & Griffin 1938; Schnitzler, Moss & Denzinger 2003), and acoustic surveys have become an increasingly popular alternative method or addition to conventional bat survey methods. Acoustic surveys can be carried out in a wide range of habitats to detect a large number of species. For example, they can detect more of the aerial insectivorous species present than netting and trapping methods (O'Farrell & Gannon 1999; MacSwiney, Clarke & Racey 2008). They can also be used in habitats where capture methods are inefficient or difficult, such as large open fields and high in the forest canopy (Kunz, Hodgkinson & Weise 2009), and allow cost-effective, long-term autonomous monitoring.

Reliable species identifications are critical for survey and monitoring programmes. Many bat species have evolved species-specific echolocation call structures, shaped by ecological and perceptual selection pressures (Fenton & Bell 1981; Jones & Teeling 2006), facilitating acoustic identification (Ahlén & Baagøe 1999). However, call structures within species can be extremely flexible and depend on factors including habitat, age, sex and the presence of conspecifics (Kalko & Schnitzler 1993; Obrist 1995; Murray, Britzke & Robbins 2001; Schnitzler, Moss & Denzinger 2003; Jones & Siemers 2011). Calls of individuals also vary depending on the sensory objectives of the bat. Whilst commuting and searching for prey, 'search-phase' calls are used, with calls becoming of a shorter duration and more rapid as the bat approaches prey (Griffin, Webster & Michael 1960; Schnitzler & Kalko 2001). Some species also exhibit call variation across their geographic range (Heller & von Helversen 1989; Siemers et al. 2005; Papadatou, Butlin & Altringham 2008; Buckley et al. 2011; Puechmaille et al. 2011b), which can complicate species identification across wide areas. Finally,

different species often exhibit overlap in call frequencies and shape, due either to convergence because of similar sensory challenges or to phylogenetic constraints (Preatoni et al. 2005; Jones & Teeling 2006).

Methods of identifying bats acoustically vary from direct assessment of a sound by listening to the output from ultrasonic detectors in the field, to applying complex statistical models to recorded call sequences. However, published studies vary considerably in the parameters they measure from calls, the degree of objectivity in their methods and the repeatability of the results they derive. For continent-wide survey and monitoring programs that aim to assess changes in activity over time or between sites, a quantitative method of identification that is objective, standardized and repeatable is essential.

A number of objective and quantitative methods have been used to identify bats acoustically, including discriminant function analysis (Zingg 1990; Vaughan, Jones & Harris 1997; Parsons & Jones 2000), support vector machines (Redgwell et al. 2009), artificial neural networks (ANN) (Parsons & Jones 2000; Redgwell et al. 2009) and synergetic pattern recognition (Obrist, Boesch & Flückiger 2004). Although these previous studies accurately classify many of the species on which they are trained and prove the concept and value of quantitative call identification, they have not been made publically accessible and are restricted to a regional (often national) level [e.g. Dadia-Lefkimi-Soufli National Park, Greece (Papadatou, Butlin & Altringham 2008); Italy (Russo & Jones 2002); UK (Parsons & Jones 2000); Switzerland (Obrist, Boesch & Flückiger 2004)]. Therefore, they cannot be used to generate comparable classifications at a continental scale. Previously, developing a continental-scale classification tool has been hampered by the lack of a suitable echolocation call reference library that represents all the species likely to be encountered, and encompasses intraspecific variation in calls and variation introduced through the use of different recording methods. The recent development of a global call library of full-spectrum calls, EchoBank (Collen 2012), paves the way for the development of continent-wide echolocation classification tools.

We develop our classifiers using reference search-phase echolocation calls from EchoBank to train ensembles of artificial neural networks (eANNs) to distinguish calls from 34 European bat species. ANNs have been used successfully to classify the vocalizations of marine mammals (Murray, Mercado & Roitblat 1998), identify acoustically conspicuous but visually concealed birds such as corncrakes *Crex crex* (Terry & McGregor 2002) and recognize species of Orthoptera (Chesmore & Ohya 2004). A number of comparisons have shown ANNs to outperform other statistical techniques in classification of echolocation calls (Parsons & Jones 2000; Redgwell et al. 2009; Armitage & Ober 2010).

We present an online and freely accessible pan-European acoustic identification tool, iBatsID, for objective and consistent identification of full-spectrum bat echolocation calls recorded throughout Europe, with a quantitative measure of uncertainty in identifications. This tool can be used for standardized continent-wide acoustic bat survey and monitoring programmes, and we demonstrate its application using data from the iBats programme (Jones et al., in press).

Materials and methods

Reference calls

Full-spectrum search-phase echolocation call sequences for 34 European bat species were obtained from EchoBank (Collen 2012) (Fig. 1a). EchoBank collates time-expanded or directly sampled call sequences from individual bats of a known species. Sequences were recorded using a variety of equipment, in different circumstances [e.g. in the hand (*Rhinolophus* species only), hand-released, light-tagged, free-flying], geographic regions (e.g. Bulgaria, France, Greece, Switzerland, UK) and habitats (open, edge, cluttered). Therefore, the data set encompasses a high degree of intraspecific call variation and flexibility (Fig. 1b and see Table S1). Only calls recorded in Europe were used, with the exception of *Eptesicus bottae* for which only calls recorded in Israel and Egypt were available. These were used to represent the *Eptesicus bottae/antolicus* complex in Europe.

European species for which reference calls were unavailable were excluded from the classification tool. Following the taxonomy of Simmons (2005) and the definition of Europe used in the IUCN European Mammal assessment (extending from Iceland to the Urals and Franz Josef Land to the Mediterranean, excluding the Caucasus region and including the Canary Islands, Madeira and the Azores) (Temple & Terry 2007), the species we did not include are four island endemic species (*Plecotus sardus*, *Plecotus teneriffae*, *Nyctalus azoreum* and *Pipistrellus maderensis*) and two mainland species (*Plecotus kolombatovici* and *Plecotus macrobullaris*). However, we acknowledge that taxonomic revisions and new species descriptions have added further species to the list since 2005, including *Pipistrellus hanaki*, *Eptesicus isabellinus*, *Myotis aurascens* and *Myotis escaleraei*. As the majority of our reference calls were classified to species prior to these taxonomic revisions, we use the earlier taxonomy here, but caution that calls identified as species such as *Myotis nattereri*, *Myotis mystacinus*, *E. bottae* and *Eptesicus serotinus* may in fact represent species complexes.

We used the commercially available sound analysis software, SonoBat version 3 (Szewczak 2010), to automatically find and measure calls in the recorded sequences. SonoBat uses amplitude threshold filters and recognition of smooth frequency changes over time to find calls and to fit a frequency–time trend line to the shape of the call, from which a number of measurements are extracted. Automatic feature extraction removes operator measurement bias from call parameters. All calls located by SonoBat were visually inspected, and calls where the measurement line did not fit the call accurately (i.e. the fitted line included background noise or echo) were rejected, using a customized accept/reject button in SonoBat (Szewczak 2010, personal communication). We measured 15 858 search-phase calls in 1259 sequences, with each sequence assumed to be from a different individual.

To have sufficient data to train and test the eANNs for each species, we set a minimum sample size of 26 calls, as this was the minimum total number of calls available for any species. Calls in each sequence were ranked by call quality (highest quality calls have the highest signal-to-noise ratio without being overloaded and without overlap between the call and an echo) and the highest quality call in each sequence was selected. To attain the minimum sample size for the 12 species with fewer than 26 sequences available, multiple calls from each sequence were chosen, with highest quality calls selected first. A total of 1350 calls were selected. Although the highest quality call in each sequence was selected, the quality of selected calls were variable (call-quality score mean = 0.89, SD = 0.10, range = 0.41–0.98 out of a possible range of 0–1), which helps to ensure that the tool can classify calls of variable quality recorded in real-world situations.

Parameter selection

Twenty-four parameters describing the frequency and time course of the call were automatically extracted by SonoBat (Table S2). Species were grouped according to echolocation call type, reflecting either phylogenetic constraints or convergence (groups 1–5, see Fig. 1a). Parameters were compared within each group to determine which are most useful in distinguishing species, as the most useful parameters may differ depending on the type of call used. We compared the variance in each parameter among species within each group to the variance within each species (F-ratio of univariate anova). As F-ratios >1 suggests that interspecific call variation is greater than intraspecific variation, we assumed that the parameters with higher F-ratios would be more useful in distinguishing between species. A k-means cluster analysis was used within each call-type group to separate parameters into two clusters based on the F-ratios, and those in the high mean cluster were selected. Parameters selected as important for any group were then used to build each stage of the neural networks, giving a total of 12 parameters (Table S2). Correlated parameters were not removed, as the extent of correlation differed between species, and these differences were deemed important in classification. Numbers of calls and sequences used for each species, along with mean parameter values, are shown in Table S3. Statistical tests were carried out in R version 2.13 (R Development Core Team 2011).

Ensembles of artificial neural networks

We used eANNs (multi-layer perceptrons) to develop iBatsID. ANNs are machine-learning methods trained to classify input data into particular output categories (Rumelhart, Hinton & Williams 1986). They contain networks of interconnected processing units (neurons) arranged into layers, with each unit connected to every unit in the preceding layer. A subset of data, for which the output categories are known, is used to train the network, altering the strength of the connections between units. During this process, networks can learn from their mistakes to maximize classification rates. An independent set of data with known classifications is used to assess the accuracy of the ANN and provide an independent confidence measure for classification to each category.

Using ensembles of ANNs achieves a higher classification rate than any single classifier as long as the classification rate for each ANN within the ensemble is greater than 50% (Redgwell et al. 2009). Here, we assembled ensembles in a hierarchical structure, to further increase classification accuracy, and to enable classification to genus or subgroup level for those calls that cannot be confidently classified to species level. The first ensemble was trained to classify to one of the five call-type groups, and subsequent ensembles were trained to classify to species level within each group. Where classification to species level yielded classification accuracy of $<70\%$ for any species, further ensembles were trained on subgroups of species within each call-type group to improve accuracy.

Ensembles of artificial neural networks were trained and tested using a custom-written Java application following the methodology in Redgwell et al. (2009). Half of the data were used to train the networks and half used as an independent testing data set to assess accuracy. eANNs used a

sigmoidal activation function. Nine permutations of learning rate, six of momentum, two of number of hidden layers and 11 of number of neurons per hidden layer were run (see Redgwell et al. 2009 for details). The top-performing 50 networks, judged based on the highest minimum classification accuracy, were retrained 20 times with randomly initialized unit connection weights as initial weights can affect classification accuracy. Of these retrained networks, the 21 top-performing networks were used as an ensemble. At each stage of the hierarchy, probability of correct classification was calculated as the product of the percentage of calls classified correctly up to and including that stage. iBatsID and instructions for use, as well as network configurations for all trained networks, are available online at <http://sites.google.com/site/ibatsresources/iBatsID>.

Thresholds

As each classifier within the ensemble votes on the classification, the output is probabilistic and a 'majority rules' system determines species identification. Therefore, it is possible that with five species, a call can be classified to a particular species with only 21% of the votes. We applied a series of thresholds to prevent ambiguous classification, at 70%, 80%, 90% and 95% such that only calls attributed to a particular call-type group/subgroup/species with this share of the votes are classified. This introduces a trade-off between maximizing correct classifications and minimizing the number of unclassified calls. The median probabilities of calls being classified correctly, misclassified and unclassified at these different threshold levels, were calculated.

Application to iBats data

To illustrate how the tool can be used, the eANNs were used to classify calls from noisy recordings (call-quality score mean = 0.81, SD = 0.17, range = 0.3–0.98) from car-based acoustic transects from Ukraine in 2009–2011, as part of the iBats project (Jones et al., in press). Thresholds of 70%, 80%, 90% and 95% were applied to eANN outputs to assess the effects on call classification rates.

Results

Parameter selection

Parameters useful for classifying species were similar within each call-type group (Fig. 2a–e), although only FPeak was useful for all groups. Overall, 12 parameters were selected, of which eight describe different aspects of the frequency of the call: FMax, FMin, BW, FCtr, FC, FPeak, FLg and FKn; one describes the duration of the call: Dur; and three describe the change in frequency over time (slope) of the call: StartS, SteepS and FMaxFKnS (see Table S2 for parameter definitions). The F-ratio values associated with parameters for the Myotis and Plecotus groups are much lower than for other groups (Fig. 2f), suggesting that within each of these groups, species calls are more similar.

Neural networks

Ensembles of artificial neural networks achieved an overall median correct classification rate of 83.7% (mean = 80.5%, range 48.7–100%) across the hierarchy for all 34 species (Fig. 3). Classification to call-type groups achieved a median rate of 97.6% correct, with varying classification rates to species level within each group (Fig. 3). Classification to species within the *Myotis* group was least accurate, with a median correct classification rate of only 60.3% (mean = 62.1%). Median correct classification for non-*Myotis* species was 90.4% (mean = 90.5%).

Group 1 – *Rhinolophus*

A median of 100% (mean = 94.9%) of calls within the *Rhinolophus* group were classified correctly to species level. *Rhinolophus ferrumequinum* and *Rhinolophus blasii* are identifiable by peak frequency alone, whereas *Rhinolophus hipposideros*, *Rhinolophus mehelyi* and *Rhinolophus euryale* exhibit some frequency overlap. Only *R. hipposideros* and *R. mehelyi* are occasionally misclassified; *R. mehelyi* was confused with *R. hipposideros* (5% of calls) and *R. euryale* (5% of calls); *R. hipposideros* was confused with *R. mehelyi* (11.6% of calls) and *R. euryale* (3.8% of calls) (Fig. 3).

Group 2 – *Pipistrellus* and *Miniopterus*

This call-type group was split into two subgroups, one containing *Pipistrellus kuhlii* and *Pipistrellus nathusii* and the other containing *Pipistrellus pipistrellus*, *Pipistrellus pygmaeus* and *Miniopterus schreibersii*. Classification to either subgroup is 97.6% accurate, and classification to species level within the *P. pipistrellus*/*P. pygmaeus*/*Mi. schreibersii* complex averages 96.4% correct, with *P. pygmaeus* occasionally misclassified as *P. pipistrellus* or *Mi. schreibersii*. Correct classification to species level in the *P. nathusii*/*P. kuhlii* complex is lower, averaging 83.8% (Fig. 3), because of the similarity in calls of these species.

Group 3 – *Nyctalus*, *Eptesicus*, *Hypsugo*, *Barbastella*, *Vespertilio* and *Tadarida*

This call-type group was split into a subgroup containing *E. bottae* and *Hypsugo savii* and another containing *E. serotinus*, *Nyctalus leisleri*, *Nyctalus noctula* and *Vespertilio murinus*. The remaining species within this call-type group were classified accurately in over 90% of cases without further subgrouping (Fig. 3).

Between 72.7% and 90.4% of calls in the *E. serotinus*/*N. leisleri*/*N. noctula*/*V. murinus* complex were classified correctly. Calls of species in this subgroup exhibit a high level of intraspecific variation, and overlapping frequencies are used by different species, which may have contributed to the difficulty in classification. *Vespertilio murinus* has the lowest classification rates in this subgroup, with calls misclassified as *N. leisleri* and *N. noctula* (Fig. 3).

Group 4 – Plecotus

The Plecotus species call-type group only contained two species (*Plecotus auritus* and *Plecotus austriacus*), and the eANN identifies 90·9% of Plecotus calls correctly to species level (Fig. 3).

Group 5 – Myotis

A median of 60·3% of calls within the Myotis call-type group were classified correctly. This is far lower than any other group, reflecting the spectral and temporal similarity of the echolocation calls of Myotis species within this call-type group. Best classification was achieved by creating three subgroups: the first contains *Myotis alcathoe* and *Myotis emarginatus*; the second contains *Myotis blythii*, *Myotis punicus* and *Myotis myotis* and the final subgroup contains *Myotis bechsteinii*, *Myotis brandtii*, *Myotis daubentonii* and *M. mystacinus*. Other species can be classified without further subgrouping (Fig. 3). *Myotis nattereri* calls were most easily classified, with 80·7% correct. Individuals in the *M. bechsteinii*/*M. brandtii*/*M. daubentonii*/*M. mystacinus* subgroup are most difficult to classify, with lower correct classification rates for these species than any others, ranging from 49·2% to 53·9%.

Thresholds

Applying thresholds to the classification probabilities resulted in slight improvements in the probability of correct classification, although results vary across species (Table S4). However, the proportion of unclassified calls increases notably as thresholds increase (Fig. 4); at a threshold of 95%, a median of only 9% of calls are misclassified to species level, but the median probability of being unable to classify a call is increased to 50·6%.

Application to iBats data

When applying a threshold level of 70%, only 4% of the 3630 iBats calls analysed were not classified, 92% were classified to subgroup level, and 68% of calls were classified to species level (Fig. 5a,b). Increasing the threshold to 90% reduces the number of calls classified to subgroup level to 82% and species level to 46%. Classification rates to species level in the Myotis group were lowest; however, the majority of calls can still be classified to a subgroup.

Discussion

iBatsID can be used to identify species recorded in acoustic bat survey and monitoring programs over Europe, to provide objective identification that is consistent, repeatable through time and

comparable across the continent. This will enable effective continental monitoring of bat activity and distribution patterns (Blumstein et al. 2011) and will improve the efficiency of standardized acoustic monitoring programs, such as iBats (Jones et al., in press). Such programmes can provide invaluable data on the status of bat populations and their responses to global change at a continental scale, providing practitioners with the information necessary for effective conservation strategy. They may also provide insight into the behaviour and conservation requirements of migratory species; something that regionally restricted programmes are unlikely to achieve.

The classification rates we present for some species are lower than those achieved in other studies. For example, our classification rates for *Myotis capaccinii* (72.3%) and *M. emarginatus* (70.1%) are lower than achieved by Papadatou, Butlin & Altringham (2008) (91.1% and 95.2%, respectively); our results for *P. kuhlii* (84.9%) and *M. capaccinii* are lower than obtained by Russo & Jones (2002) (98% and 88% respectively) and classification rates for all *Myotis* species, *Barbastella barbastellus* and *E. serotinus* are lower than those reported by Redgwell et al. (2009). This is almost certainly the result of our eANN dealing with many more species, which increases the overlap in parameter space between species. For example, we include 12 *Myotis* species compared to between five and nine in other studies (Parsons & Jones 2000; Obrist, Boesch & Flückiger 2004). We also include a range of geographic (habitat and regional) and methodological variation in the reference call data set, increasing the variability in calls that makes classification more difficult. However, this decrease in classification accuracy leads to a positive trade-off with the increased generality of use for the tool across Europe. This is necessary for a realistic pan-European classification tool that can provide consistent classification rates across space and time; an essential step for continent-wide bat monitoring.

Region-specific classifiers such as in Obrist, Boesch & Flückiger (2004), Papadatou, Butlin & Altringham (2008), Redgwell et al. (2009) and Russo & Jones (2002) provide a higher level of accuracy for the species most commonly found in that particular area, but regional classifiers will not necessarily produce classification results that are comparable with each other, adding a source of error to any continental monitoring effort. Also, as regional classifiers have been trained on a restricted number of species, they will not correctly identify species that may move into new areas as a result of changing climatic conditions. In the light of the predicted distribution shifts for many species under climate change (Parmesan & Yohe 2003), including European bats (Rebelo, Tarroso & Jones 2009), the ability to detect species shifting distribution patterns is fundamental for future conservation planning.

We demonstrate median correct classification rates of 60.3% for *Myotis* species, which may not be sufficient for accurate monitoring and generation of distribution data for these species. By increasing the threshold level to 90%, we can increase median classification across the 12 *Myotis* species to 73.3%. However, accepting this level of identification error may impact on our estimates of species distributions and distribution change over time. Given the difficulties in achieving good classification rates for *Myotis* species here, and elsewhere, it may be necessary to conclude that confident species-level classification in this genus is not possible with current methods. Monitoring studies should bear this in mind and could, for example, monitor *Myotis* species as a group using acoustic methods, with this tool offering a simple, objective way for bat surveyors to identify *Myotis*. Other

survey methods could then be employed for specific species which may be of interest (e.g. standardized maternity roost counts for free hanging *M. myotis* and *M. blythii* or counts in hibernacula for *Myotis dasycneme*).

Weighting methods may be applied to improve species classification locally. For example, the classification probabilities of species that are known not to occur in an area could be down-weighted relative to the distance to their known distribution, to avoid misclassification as these species. When we trialled such a system with this data set, it was found to reduce the ability to correctly classify the down-weighted species if they were to move into new areas, which would reduce our ability to identify species distributional changes. However, applying buffer zones around the current distribution, within which down-weighting is not applied, or using fuzzy logic to incorporate ecological factors and distribution data into a weighting system, may enable increased classification rates for some species, whilst not impeding our ability to detect distribution shifts. Further exploration of such weighting methods could prove useful in improving classification rates.

The ability to generate accurate species-level classification seems to be far more dependent on the call characteristics of the species involved and, in particular, the extent of call similarity between species, than on the quantity or quality of training data. This suggests that potential for acoustic identification in new areas could be assessed by modelling the extent of call similarity within species assemblages. Even in areas with low species call similarity, a good training data set may require a considerable number of call sequences for widely distributed species, to capture the full extent of their call repertoire. Here we have used minimal numbers of calls from each individual bat to avoid problems of pseudo-replication. Further investigation into the effect of using entire sequences of calls in classification may enable a far greater volume of data to be used from each recording, encompassing more of the variation in calls and making collection of suitable training data far easier.

Reference calls and parameter selection

The validity and efficacy of this classification tool are reliant on the quality of the data used to train the eANN. For most species, our call library contains recordings from a variety of methods and surroundings, providing some confidence that intraspecific variation is represented in the calls used to train the eANN. However, five species (*R. euryale*, *R. mehelyi*, *R. blasii*, *M. punicus* and *M. dasycneme*) were only recorded in one area, a limited variety of surroundings, or only a limited number of individuals or sequences were recorded so the variation captured for these species is unlikely to represent the full repertoire of call variation present over their geographic range (Russo et al. 2007). Further collection of reference calls for these species, and incorporation into the neural network of reference calls from further geographic areas for all species, will help ensure that classification is robust in as many different regions and environments in Europe as possible. At present, including further species or calls would necessitate retraining the eANNs to incorporate the extra data. Further work will focus on automating the retraining process so that new library calls can be automatically included into an updated classifier.

We have presented an objective method for selecting parameters for species classification and found 12 parameters to be most useful in distinguishing between species. These parameters include

most of the parameters used in previous classification studies (Zingg 1990; Parsons & Jones 2000; Russo & Jones 2002; Papadatou, Butlin & Altringham 2008; Redgwell et al. 2009) with five new parameters also selected here: FKn, FLg and the three slope parameters (Table S2). However, other methods of characterizing calls may be of equal or more use in classification. Reducing the total signal content by taking any number of measurements from calls results in a loss of information, and this information may be useful in increasing classification rates. Methods that consider the call in its entirety, such as synergetic image classification tools (Obrist, Boesch & Flückiger 2004) or morphometric techniques (Macleod 2001) may prove more successful in distinguishing between echolocation calls and in classifying *Myotis* species, which have very similar calls and remain problematic to classify using parameter characterization methods (Lundy et al. 2011).

Application of the tool

Preliminary analysis of iBats data suggests this tool can be used to classify calls from noisy, real-world situations, with good levels of success for all European genera except *Myotis*. At present, this identification tool should be used in conjunction with SonoBat software to automatically extract call parameters from recordings, to ensure comparability between the call parameters used to train the eANN and those measured from unclassified calls. To ensure the best opportunity for correct classification, we suggest using the best quality calls within a recorded sequence. Classifying more than one call from a sequence will help to validate the species classification. As the median classification rate across all 34 species is 83.7%, it must be expected that some calls will be misclassified. We suggest close inspection of any calls that are classified as a species not previously recorded to occur in an area, particularly if the known distribution of that species is far from the study area. Calls identified as species that could have feasibly expanded their range to include the study area should generate incentive to undertake trapping or netting effort to ascertain with a higher confidence whether the species has changed its distribution.

Conclusions

The application of iBatsID allows efficient, objective and quantitative classification of potentially huge continent-scale bat survey results. While it may not provide sufficient confidence to classify unknowns to species level within all call-type groups, it is a useful means of monitoring over wide geographical scales, to develop bioindicators, detect distribution changes for many European bats and to inform conservation decisions.

Acknowledgements

This work was financially supported by a NERC studentship to C.L.W (NE/H525003/1), a sabbatical to S.P. from The University of Auckland, a Leverhulme Trust award to K.E.J and The Darwin Initiative (15033, EIDPO036, EIDPR075). We also thank Nancy Jennings (<http://dotmoth.co.uk>) and Michel Barataud for providing echolocation calls, Jon Russ, Andriy-Taras Bashta and the Ukraine iBats volunteers for data collection, Catherine Sayer, Joe Szewczak and Victor Obolonkin for assistance and Danilo Russo and two anonymous referees for helpful comments on the manuscript. This study is dedicated to the memory of Björn Siemers, who passed away before publication. We miss him, his enthusiasm for life and his massive contribution to our understanding of bat acoustics and evolution.

References

- Ahlén, I. & Baagøe, H. (1999) Use of ultrasound detectors for bat studies in Europe: experiences from field identification, surveys, and monitoring. *Acta Chiropterologica*, 1, 137–150.
- Armitage, D.W. & Ober, H.K. (2010) A comparison of supervised learning techniques in the classification of bat echolocation calls. *Ecological Informatics*, 5, 465–473.
- Battersby, J. (Comp) (2010) Guidelines for Surveillance and Monitoring of European Bats. EUROBATS publication series No. 5. UNEP/EUROBATS Secretariat, Bonn, Germany.
- Blumstein, D., Mennill, D., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J., Krakauer, A., Clark, C., Cortopassi, K., Hanser, S., McCowan, B., Ali, A. & Kirschel, A. (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48, 758–767.
- Buckley, D., Puechmaille, S., Roche, N. & Teeling, E. (2011) A critical assessment of the presence of *Barbastella barbastellus* and *Nyctalus noctula* in Ireland with a description of *N. leisleri* echolocation calls from Ireland. *Hystrix Italian Journal of Mammalogy*, 22, 111–127.
- Chesmore, E.D. & Ohya, E. (2004) Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. *Bulletin of Entomological Research*, 94, 319–330.
- Collen, A. (2012) The evolution of echolocation. PhD thesis, University College London, London, UK.
- Fenton, M. & Bell, G. (1981) Recognition of species of insectivorous bats by their echolocation calls. *Journal of Mammalogy*, 62, 233–243.
- Griffin, D.R., Webster, F.A. & Michael, C.R. (1960) The echolocation of flying insects by bats. *Animal Behaviour*, 8, 141–154.
- Heller, K. & von Helversen, O. (1989) Resource partitioning of sonar frequency bands in rhinolophoid bats. *Oecologia*, 80, 178–186.
- Jones, G. & Siemers, B.M. (2011) The communicative potential of bat echolocation pulses. *Journal of Comparative Physiology A*, 197, 447–457.
- Jones, G. & Teeling, E. (2006) The evolution of echolocation in bats. *Trends in Ecology & Evolution*, 21, 149–156.
- Jones, G., Jacobs, D., Kunz, T., Willig, M. & Racey, P. (2009) Carpe noctem: the importance of bats as bioindicators. *Endangered Species Research*, 8, 93–115.
- Jones, K.E. & Maclarnon, A. (2001) Bat life histories: testing models of mammalian life-history evolution. *Evolutionary Ecology*, 3, 465–476.
- Jones, K.E., Russ, J., Bashta, A.-T., Bilhari, Z., Catto, C., Csősz, I., Gorbachev, A., Győrfi, P., Hughes, A., Ivashkiv, I., Koryagina, N., Kurali, A., Langton, S., Maltby, A., Margiean, G., Pandourski, I., Parsons, S., Prokofev, I., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Walters, C., Weatherill, A. & Zavarzin, O (in press) Indicator Bats Program: a system for the global acoustic monitoring of bats. *Biodiversity Monitoring and Conservation: Bridging the Gaps Between Global Commitment and Local Action* (eds B. Collen, N. Pettorelli, S. Durant, L. Krueger & J. Baillie) Wiley-Blackwell, London.

Kalko, E.V. & Schnitzler, H.-U. (1993) Plasticity in echolocation signals of European Pipistrelle bats in search flight: implications for habitat use and prey detection. *Behavioral Ecology and Sociobiology*, 33, 415–428.

Kery, M. & Schmid, H. (2005) Estimating species richness: calibrating a large avian monitoring programme. *Journal of Applied Ecology*, 43, 101–110.

Kunz, T.H., Hodgkinson, R. & Weise, C. (2009) Methods of capturing and handling bats. *Ecological and Behavioural Methods for the Study of Bats*, 2nd edn (eds T.H. Kunz & S. Parsons), pp. 3–35. Johns Hopkins University Press, Baltimore.

Lundy, M., Teeling, E., Boston, E., Scott, D., Buckley, D., Prodohl, P., Marnell, F. & Montgomery, I. (2011) The shape of sound: elliptic fourier descriptors (EFD) discriminate the echolocation calls of *Myotis* bats (*M. daubentonii*, *M. nattereri* & *M. mystacinus*). *Bioacoustics*, 20, 101–116.

Macleod, N. (2001) Landmarks, localization and the use of morphometrics in phylogenetic analysis. *Fossils, Phylogeny and Form: An Analytical Approach* (eds G. Edgecombe, J. Adrain & B. Lieberman), pp. 197–234. Plenum, New York.

MacSwiney, M.C., Clarke, F.M. & Racey, P. (2008) What you see is not what you get: the role of ultrasonic detectors in increasing inventory completeness in Neotropical bat assemblages. *Journal of Applied Ecology*, 45, 1364–1371.

Murray, K., Britzke, E. & Robbins, L. (2001) Variation in search phase calls of bats. *Journal of Mammalogy*, 82, 728–737.

Murray, S.O., Mercado, E. & Roitblat, H.L. (1998) The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. *Journal of the Acoustic Society of America*, 104, 3626–3633.

Obrist, M.K. (1995) Flexible bat echolocation: the influence of individual, habitat and conspecifics on sonar signal design. *Behavioral Ecology and Sociobiology*, 36, 207–219.

Obrist, M.K., Boesch, R. & Flückiger, P.F. (2004) Variability in echolocation call design of 26 Swiss bat species: consequences, limits and options for automated field identification with a synergetic pattern recognition approach. *Mammalia*, 68, 307–322.

Ochoa, J.G., O'Farrell, M.J. & Miller, B.W. (2000) Contribution of acoustic methods to the study of insectivorous bat diversity in protected areas from northern Venezuela. *Acta Chiropterologica*, 2, 171–183.

O'Farrell, M.J. & Gannon, W.L. (1999) A comparison of acoustic versus capture techniques for the inventory of bats. *Journal of Mammalogy*, 80, 24–30.

Papadatou, E., Butlin, R.K. & Altringham, J.D. (2008) Identification of bat species in Greece from their echolocation calls. *Acta Chiropterologica*, 10, 127–143.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421, 37–42.

Parsons, S. & Jones, G. (2000) Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *The Journal of Experimental Biology*, 203, 2641–2656.

Pereira, H.M. & Cooper, H.D. (2006) Towards the global monitoring of biodiversity change. *Trends in Ecology & Evolution*, 21, 123–129.

Pierce, G. & Griffin, D. (1938) Experimental determination of supersonic notes emitted by bats. *Journal of Mammalogy*, 19, 454–455.

Poiani, K.A., Richter, B., Anderson, M. & Richter, H. (2000) Biodiversity conservation at multiple scales: functional sites, landscapes, and networks. *BioScience*, 50, 133–146.

Preatoni, D.G., Nodari, M., Chirichella, R., Tosi, G., Wauters, L.A. & Martinoli, A. (2005) Identifying bats from time-expanded recordings of search calls: comparing classification methods. *The Journal of Wildlife Management*, 69, 1601–1614.

Puechmaille, S.J., Frick, W.F., Kunz, T.H., Racey, P.A., Voigt, C.C., Wibbelt, G. & Teeling, E.C. (2011a) White-nose syndrome: is this emerging disease a threat to European bats? *Trends in Ecology & Evolution*, 26, 573–579.

Puechmaille, S.J., Gouilh, M.A., Piyapan, P., Yokubol, M., Mie, K.M., Bates, P.J., Satasook, C., Nwe, T., Bu, S.S.H., Mackie, I.J., Petit, E.J. & Teeling, E.C. (2011b) The evolution of sensory divergence in the context of limited gene flow in the bumblebee bat. *Nature Communications*, 2, 573.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org>.

Rebelo, H., Tarroso, P. & Jones, G. (2009) Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology*, 16, 561–576.

Redgwell, R.D., Szewczak, J.M., Jones, G. & Parsons, S. (2009) Classification of echolocation calls from 14 species of bat by support vector machines and ensembles of neural networks. *Algorithms*, 2, 907–924.

Robbins, C.S., Bystrak, D. & Geissler, P.H. (1986) The Breeding Bird Survey: its first fifteen years, 1965–1979. US Fisheries and Wildlife Service Resource Publication 157.

Rumelhart, D., Hinton, G. & Williams, R. (1986) Learning representations by back-propagating errors. *Nature*, 323, 533–536.

Russo, D. & Jones, G. (2002) Identification of twenty-two bat species (Mammalia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls. *Journal of Zoology*, 258, 91–103.

Russo, D., Mucedda, M., Bello, M., Biscardi, S., Pidinchiedda, E. & Jones, G. (2007) Divergent echolocation call frequencies in insular rhinolophids (Chiroptera): a case of character displacement? *Journal of Biogeography*, 34, 2129–2138.

Schnitzler, H. & Kalko, E.K.V. (2001) Echolocation by insect-eating bats. *BioScience*, 51, 557–569.

Schnitzler, H.-U., Moss, C.F. & Denzinger, A. (2003) From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology & Evolution*, 18, 386–394.

Siemers, B.M., Beedholm, K., Dietz, C., Dietz, I. & Ivanova, T. (2005) Is species identity, sex, age or individual quality conveyed by echolocation call frequency in European horseshoe bats? *Acta Chiropterologica*, 7, 259–274.

Simmons, N.B. (2005) Order Chiroptera. *Mammal Species of the World: A Taxonomic and Geographic Reference* (eds D.E. Wilson & D.M. Reeder), pp. 312–529. John Hopkins University Press, Baltimore.

Szewczak, J.M. (2010) SonoBat v.3. www.sonobat.com.

Temple, H.J. & Terry, A. (2007) *The Status and Distribution of European Mammals*. IUCN, Luxembourg: Office for Official Publications of the European Communities.

Terry, A. & McGregor, P. (2002) Census and monitoring based on individually identifiable vocalizations: the role of neural networks. *Animal Conservation*, 5, 103–111.

Vaughan, N., Jones, G. & Harris, S. (1997) Identification of British bat species by multivariate analysis of echolocation call parameters. *Bioacoustics*, 7, 189–207.

Zingg, P.E. (1990) Akustische Artidentifikation von Fledermäusen (Mammalia: Chiroptera) in der Schweiz [Acoustic species identification of bats (Mammalia, Chiroptera) in Switzerland]. *Revue Suisse de Zoologie*, 97, 263–294.

Figures

Figure 1.

Spectrograms of representative search-phase echolocation calls (Hanning window and FFT size of 512) showing (a) inter-specific variability for 34 species separated into call-type groups (1 – *Rhinolophus*; 2 – *Pipistrellus*, *Miniopterus*; 3 – *Barbastella*, *Eptesicus*, *Hypsugo*, *Nyctalus*, *Tadarida*, *Vespertilio*; 4 – *Plecotus*; 5 – *Myotis*), and (b) intra-specific variability in an example species – *Nyctalus leisleri*.

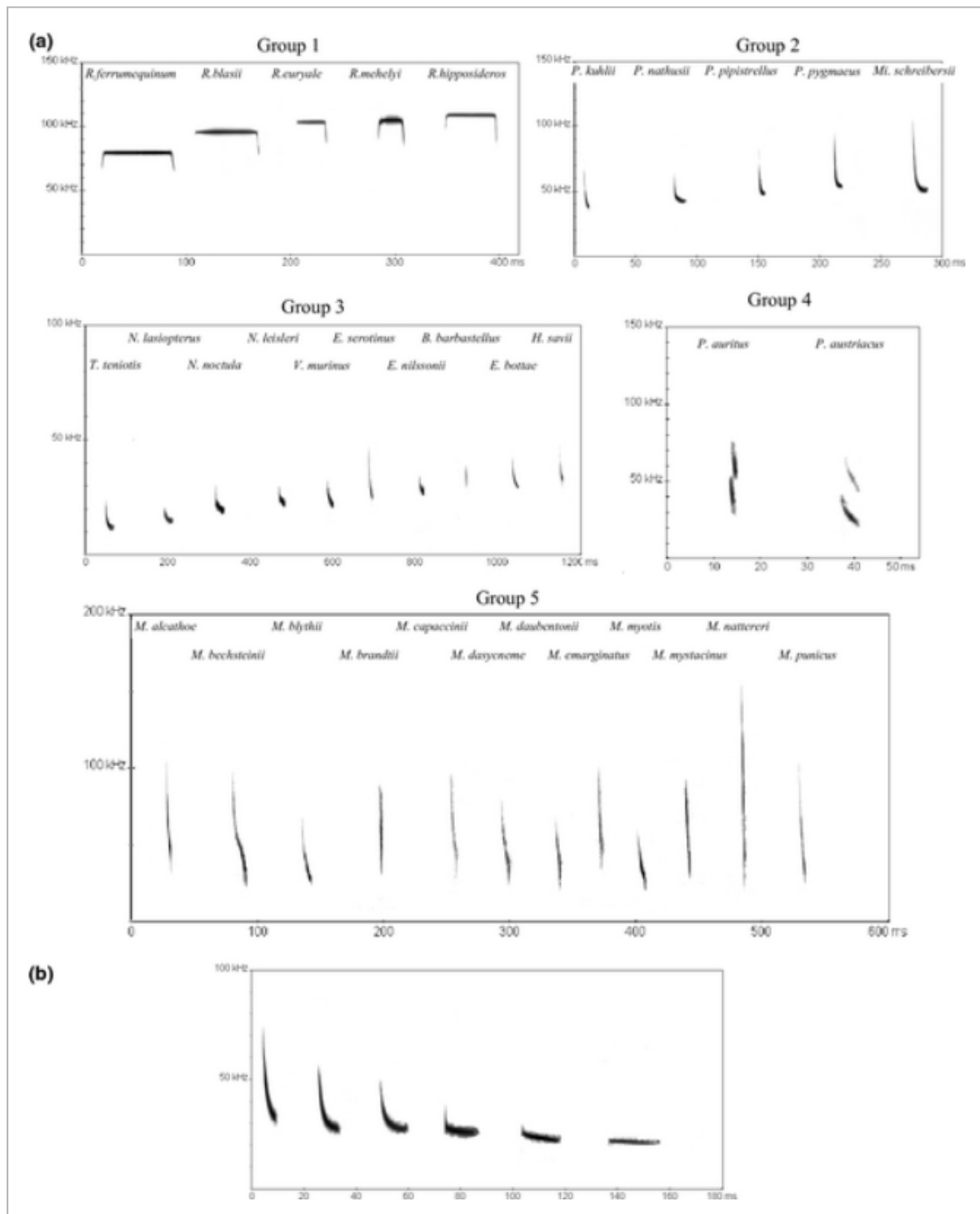


Figure 2.

F-ratios for candidate parameters for training the ensembles of artificial neural networks, showing (a–e) Normalized F-ratios for each parameter within each call-type group. Shaded parameters are those selected for each call-type group. (f) Average F-ratios of the 12 overall selected parameters for each call-type group. Call-type groups: 1 – *Rhinolophus*; 2 – *Pipistrellus*, *Miniopterus*; 3 – *Barbastella*, *Eptesicus*, *Hypsugo*, *Nyctalus*, *Tadarida*, *Vespertilio*; 4 – *Plecotus*; 5 – *Myotis*. Parameters are as described in Table S2.

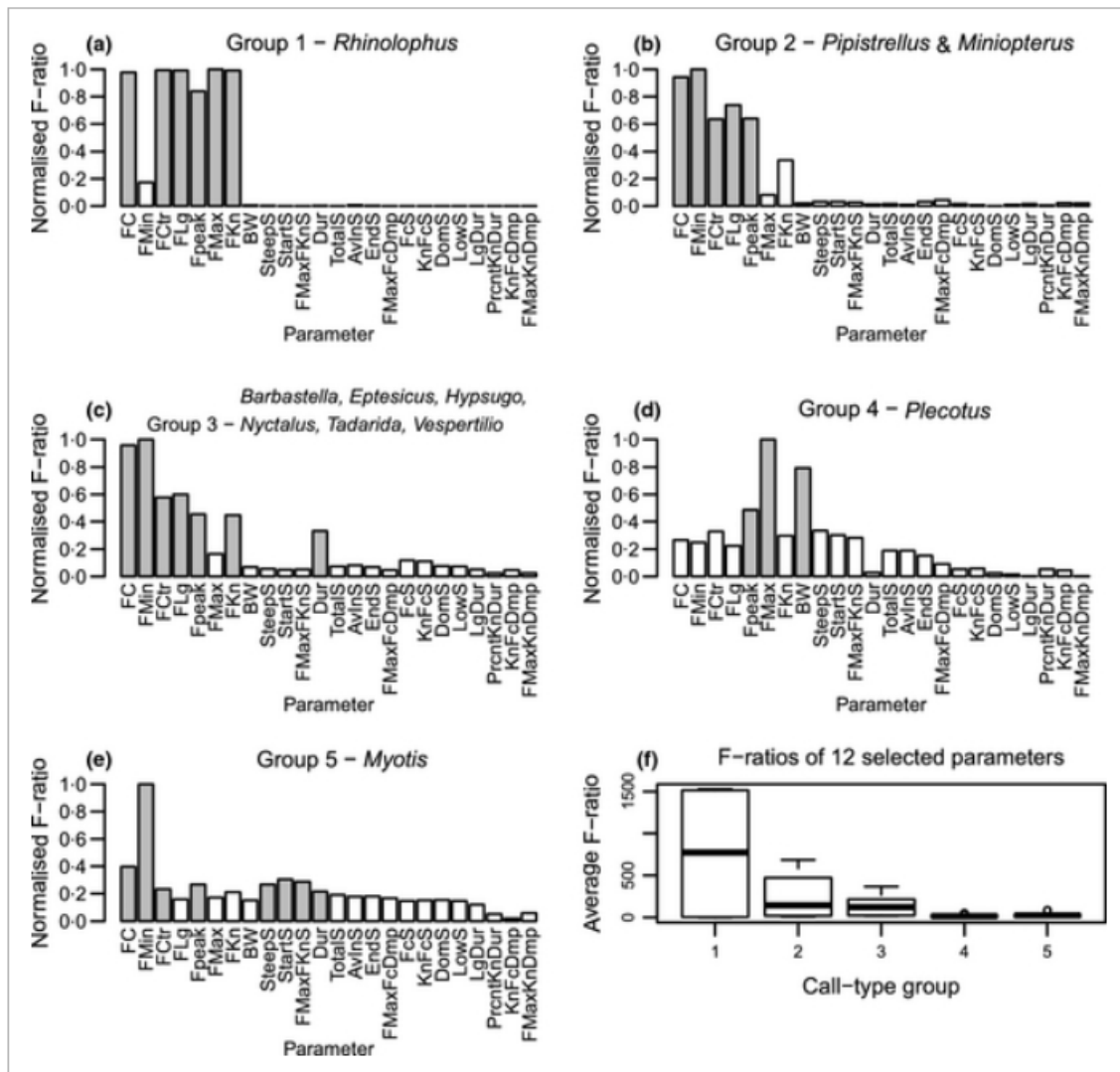


Figure 3.

Hierarchy of the ensembles of artificial neural network with percentage of testing calls classified correctly below each branch, and the percentage of false-positive results attributed to each class in parenthesis. The probability of correctly classifying calls (%) at each stage, calculated as the product of the percentage of calls classified correctly up to and including that stage, is given in bold above each branch. Probability of correct species-level classification (%) is shown on the right. Bbar, *Barbastella barbastellus*; Ebot, *Eptesicus bottae*; Enil, *Eptesicus nilssonii*; Eser, *Eptesicus serotinus*; Hsav, *Hypsugo savii*; Msch, *Miniopterus schreibersii*; Malc, *Myotis alcathoe*; Mbec, *Myotis bechsteinii*; Mbly, *Myotis blythii*; Mbra, *Myotis brandtii*; Mcap, *Myotis capaccinii*; Mdas, *Myotis*

dasycneme; Mdau, *Myotis daubentonii*; Mema, *Myotis emarginatus*; Mmyo, *Myotis myotis*; Mmys, *Myotis mystacinus*; Mnat, *Myotis nattereri*; Mpun, *Myotis punicus*; Nlas, *Nyctalus lasiopterus*; Nlei, *Nyctalus leisleri*; Nnoc, *Nyctalus noctula*; Pkuh, *Pipistrellus kuhlii*; Pnat, *Pipistrellus nathusii*; Ppip, *Pipistrellus pipistrellus*; Ppyg, *Pipistrellus pygmaeus*; Paur, *Plecotus auritus*; Paus, *Plecotus austriacus*; Rbla, *Rhinolophus blasii*; Reur, *Rhinolophus euryale*; Rfer, *Rhinolophus ferrumequinum*; Rhip, *Rhinolophus hipposideros*; Rmeh, *Rhinolophus mehelyi*; Tten, *Tadarida teniotis*; Vmur, *Vespertilio murinus*.

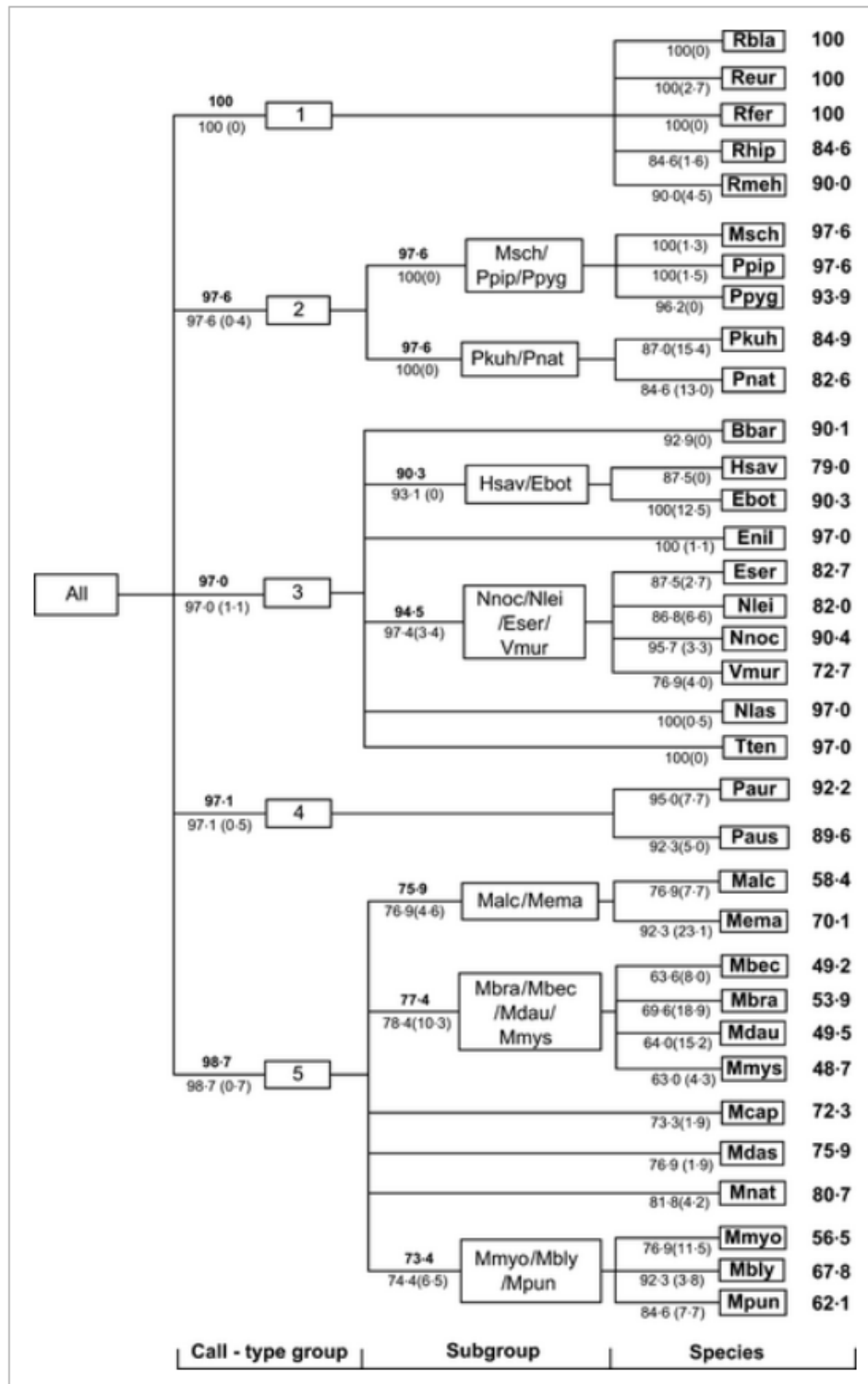


Figure 4.

The median percentage of calls classified correctly, misclassified and unclassified, at different threshold levels, averaged across the hierarchy.

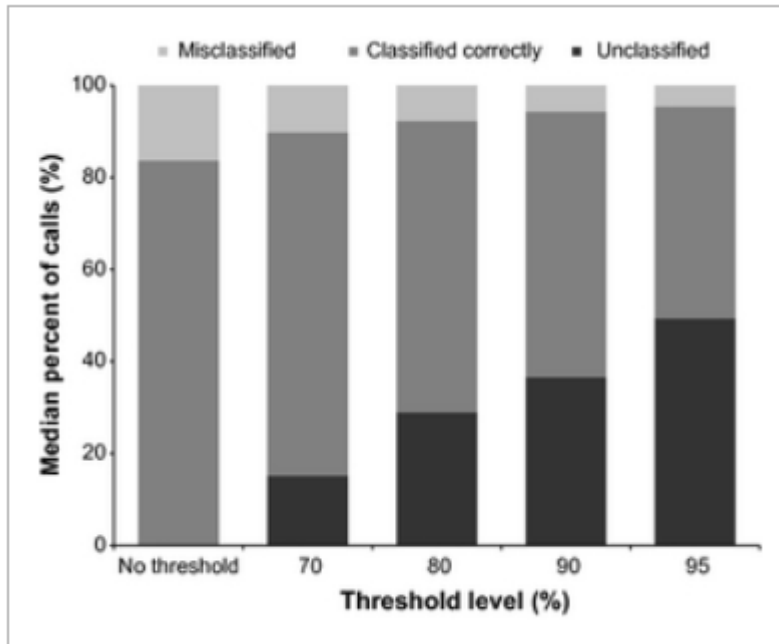


Figure 5

Application of iBatsID to data from the iBats Programme. (a) A spectrogram of an example call and classification probabilities given at each stage of the hierarchy. (b) Percentage of calls classified to call-type group, subgroup and species level, at different thresholds, within each call-type group. Group 2 – *Pipistrellus*, *Miniopterus*; Group 3 – *Barbastella*, *Eptesicus*, *Hypsugo*, *Nyctalus*, *Tadarida*, *Vespertilio*; Group 4 – *Plecotus*; Group 5 – *Myotis*. No group 1-type calls were assessed.

